

# Two-player games : a reduction

Nicolas Vieille

December 12, 2001

The goal of this chapter, together with the next one, is to give a detailed overview of the proof of the following result.

**Theorem 1** *Let  $\Gamma$  be a two-player stochastic game with finite action and states sets. The game  $\Gamma$  has a uniform equilibrium payoff.*

The complete proof is to be found in Vieille [4],[5]. The general idea of the proof is to define a class of sets, the *solvable* sets, which can safely be thought of as absorbing states, and to construct  $\varepsilon$ -equilibrium profiles such that the induced play reaches one of these sets with high probability, and remains there.

Solvable sets are defined in Section 1. The organization of the proof is presented in Section 2. We shall use the tools introduced in a chapter by Solan [2].

## 1 Solvable states

Let  $(\alpha, \beta)$  be a stationary profile, and  $C \subseteq S$  be a weakly communicating set for  $(\alpha, \beta)$ . Denote by  $\mathcal{R}$  the collection of ergodic sets for  $(\alpha, \beta)$ , which are subsets of  $C$ , and, for  $R \in \mathcal{R}$ , set  $\gamma(R, \alpha, \beta) = \lim_n \gamma_n(z, \alpha, \beta)$ , where  $z \in R$ . The limit exists and is independent of  $z$ .

**Definition 2**  *$(C, (\alpha, \beta))$  is solvable if, for some distribution  $\mu$  on  $\mathcal{R}$ , one has, for every  $z \in C$  and  $a \in A$*

$$\bigwedge_{R \in \mathcal{R}} \mu(R) \gamma^1(R, \alpha, \beta) \geq E \bigwedge v^1 | z, a, \beta_z \quad \square$$

*and a symmetric property for player 2.*

This concept is a slight generalization of the concept of *easy initial states*, introduced by Vrieze and Thuijsman.

Here,  $\beta_z$  stands for the  $z$ -component of  $\beta$ . The quantity  $\sum_{R \in \mathcal{R}} \mu(R) \gamma(R, \alpha, \beta)$  is called the solvable payoff on  $C$ . In words, there is one convex combination of the average payoffs on ergodic sets, that is individually rational, in the sense that it is at least the expected level of punishment, given any one-stage deviation of either player.

When no confusion may arise, we will omit  $(\alpha, \beta)$ , and speak of solvable sets.

Solvable sets do exist (see Thuijsman and Vrieze [3] or Vieille [?] for two players, Vieille [6] for a generalization to  $N$  player-games).

**Lemma 3** *Let  $C$  be solvable. The solvable payoff on  $C$  is an equilibrium payoff, when the initial state belongs to  $C$ .*

Corresponding  $\varepsilon$ -equilibrium strategies are designed as follows:  $(\alpha, \beta)$  is perturbed (in an history-dependent way) in such a way that the induced play remains in  $C$ , visits infinitely many times each  $R \in \mathcal{R}$ , and the average payoff  $\gamma_n(z_1, \alpha, \beta)$  is close to the solvable payoff. This profile is sustained by threats.

**Remark 4** *for the existence of solvable states, and for this lemma to hold, the assumption of perfect monitoring can be weakened. One may only assume that states and payoffs are known to the players*

Replace each state  $z$  which belongs to some solvable set  $C$  by an absorbing state, which receives as payoff the solvable payoff of  $C$  (which specific payoff is chosen for a state which belongs to several solvable sets is of no importance). It is not difficult to check that, for the *simplified* game, solvable sets coincide with absorbing states. Moreover, equilibrium payoffs of the *simplified* game are equilibrium payoffs of the *original* game.

Therefore, since we deal with the *existence* issue, we might and do assume that the only solvable sets of the game we are dealing with are absorbing states.

We denote by  $S^* \subseteq S$  the subset of non-absorbing states. For simplicity, we assume  $r^1(\cdot) < 0 < r^2(\cdot)$ . An  $\varepsilon$ -equilibrium profile  $(\sigma, \tau)$  is *absorbing* if the probability of reaching an absorbing state in finite time is at least  $1 - \varepsilon$ , whatever be the initial state.

## 2 Overview

Let  $\Gamma$  be a game. We start with a preliminary observation. Let  $C \subseteq S^*$ , and  $Q$  an exit distribution from  $C$ , that is controllable for *any* payoff vector  $\gamma \geq v$  (we refer to such a pair as a *controlled set*). Replace  $C$  by a dummy state, in which transitions are given by  $Q$  (and payoffs are arbitrary). The resulting game  $\Gamma_C$  is called *reduced*. A crucial consequence of the controllability notion is that, if  $\Gamma_C$  has an absorbing  $\varepsilon$ -equilibrium for each  $\varepsilon$ , so does  $\Gamma$ . A similar reduction can be done for a family  $\mathcal{C}$  of disjoint controlled sets. To avoid confusion, all objects related to  $\Gamma_C$  are indexed by  $\mathcal{C}$ .

We now fix the agenda. We first construct a (possibly empty) family of disjoint controlled sets  $\mathcal{C}$ , such that the reduced game  $\Gamma_C$  is nice, in a sense to be made precise. The properties of  $\Gamma_C$  enables one to construct an auxiliary recursive game (with specific features  $F$ ), such that: any absorbing  $\varepsilon$ -equilibrium profile of this auxiliary recursive game is an absorbing  $\varepsilon$ -equilibrium profile of the reduced game (provided one adds one simple threat). The final step is the proof that recursive games with features  $F$  do have absorbing  $\varepsilon$ -equilibrium profiles.

Let us be more specific.

**Definition 5** A pair  $(\beta, D)$ , where  $D \subseteq S^*$ , is a *blocking pair for player 1* if for each  $z \in D, a \in A$ ,

$$p(D|s, a, \beta_z) < 1 \Rightarrow E v^1|s, a, \beta_z < \max_D v^1$$

We also define blocking pairs  $(\alpha, D)$  for player 2 by exchanging the roles of the two players. We extend this definition to games obtained by reducing  $\Gamma$ .

**Definition 6** Let  $(C_1, \dots, C_M)$  be disjoint controlled sets. A pair  $(\beta, D)$ , where  $D \subseteq S_C^*$ , is a *reduced blocking pair for player 1* if for each  $z \in D, a \in A$ ,

$$p_C(D|s, a, \beta_z) < 1 \Rightarrow E_C v^1|s, a, \beta_z < \max_D v^1$$

We stress the fact that the value  $v^1$  that is used is the value associated with the *original* game, and *not* the value  $v_C^1$  of the reduced game.<sup>1</sup> There is no relation between reduced blocking pairs and blocking pairs of the reduced game.

---

<sup>1</sup>There is no specific relation between  $v$  and  $v_C$ . In particular,  $v_C$  may depend on the choice of the payoff in the dummy states which replace the controlled sets.

**Remark 7** *It is not crucial to define reduced blocking pairs using  $E_{\mathcal{C}}$  rather than  $E$ . Our results and proofs would remain valid if Definition 6 was modified accordingly.*

**Remark 8** *The set  $S_{\mathcal{C}}^*$  which appears in Definition 6 is the state space of the reduced game  $\Gamma_{\mathcal{C}}$ . For simplicity, we shall not distinguish between  $S_{\mathcal{C}}^*$  and  $S^*$ , in that we shall identify any subset of  $S_{\mathcal{C}}^*$  to the corresponding subset of  $S^*$ . One may wonder why reduced games are not defined by replacing each state in a given controlled set by a different state. The reason is the following. Let  $D$  be a controlled set, and replace each state in  $D$  by a dummy state as suggested. Then it may well be the case that communicating sets exist in the resulting game, which contain some but not all of the states of  $D$ . This would create many small complications. This phenomenon does of course not arise with our definition.*

Propositions 9, 10 and 11 below imply Theorem 1

**Proposition 9** *There is a collection  $\mathcal{C}$  of disjoint controlled sets such that there is no reduced blocking pair for player 2.*

Of course, there is nothing specific about player 2. The same result holds with player 1 instead.

Let  $\mathfrak{P}$  be the game obtained from  $\Gamma_{\mathcal{C}}$  by setting the payoff function in non-absorbing states to zero.

The game  $\mathfrak{P}$  is recursive and has the following features  $F$ :

- all absorbing payoffs of player 2 (resp. of player 1.) are positive (resp. negative).
- for every  $\alpha$ , there exists  $\beta$ , such that  $(\alpha, \beta)$  is absorbing.

**Proposition 10** *If  $\mathfrak{P}$  has an equilibrium payoff, the game  $\Gamma_{\mathcal{C}}$  has an absorbing  $\varepsilon$ -equilibrium profile, for every  $\varepsilon$ .*

**Proposition 11** *Every recursive game with features  $F$  has an equilibrium payoff.*

The proof of Proposition 11 is presented in the next chapter.

### 3 The reduction

We prove Proposition 10. The construction of controlled sets is done in two steps. It is elementary, except for one point. We give the idea, discuss in more detail the delicate issue, and briefly explain how to implement the idea.

#### 3.1 The idea

We start with a crucial observation. Denote by  $\alpha_\lambda$  an optimal strategy of player 1 in the  $\lambda$ -discounted zero-sum game associated with  $\Gamma$  :

$$\gamma_\lambda^1(z, \alpha_\lambda, \beta) \geq v_\lambda^1(z)$$

for every  $z, \beta$ . Denote by  $\bar{\alpha} = \lim_{\lambda \rightarrow 0} \alpha_\lambda$  (the limit is taken up to a subsequence). Define  $\bar{\beta}$  accordingly. The only property we use in this section is:

$$E \stackrel{\text{f}}{=} v^1|z, \bar{\alpha}_z, \beta_z \stackrel{\text{a}}{\geq} v^1(z), \text{ for every } z, \beta.$$

**Lemma 12** *Let  $(\alpha, D)$  be a blocking pair for player 2. There exists  $\bar{D} \subseteq D$ , such that: (i)  $v^2$  is constant on  $\bar{D}$ ; (ii)  $\bar{D}$  is communicating under  $(\alpha, \bar{\beta})$ ; (iii)  $(\alpha, \bar{D})$  is a blocking pair for player 2.*

**Proof.** Define  $\mathfrak{D}$  to be those states in  $D$  where  $v^2$  is maximal. Clearly,  $(\alpha, \mathfrak{D})$  is a blocking pair for player 2. In particular,  $\mathfrak{D}$  is stable for  $(\alpha, \bar{\beta})$ . Consider the subsets of  $\mathfrak{D}$  which are maximal for the (weak) communication property. At least one of them will do. ■

The previous lemma gives the clue to the reduction algorithm. Given any blocking pair  $(\alpha, D)$  for player 2, take  $\bar{D}$  as in the previous lemma.

- If it is the case that  $p(\bar{D}|z, a, \bar{\beta}) < 1$  and  $E(v^1|z, a, \bar{\beta}) \geq \max_{\bar{D}} v^1$ , for some  $z \in \bar{D}, a \in A$ , then  $\bar{D}$  is a controlled set. Indeed, choose among those pairs a pair  $(z^*, a^*)$  which maximizes  $E(v^1|z, a, \bar{\beta})$ . The exit distribution  $p(\cdot|z^*, a^*, \bar{\beta})$  from  $\bar{D}$  is controllable (for every continuation payoff vector  $\gamma \geq v$ ).
- Otherwise,  $(\bar{\beta}, \bar{D})$  is a blocking pair for player 1. One may then repeat the previous argument, with the two players exchanged. Take a set  $\overline{\bar{D}} \subseteq \bar{D}$ , as given by the previous lemma (where the two players are exchanged). Therefore,

- either the exit distribution  $p(\cdot|z^*, \bar{\alpha}, b^*)$  from  $\bar{\bar{D}}$  is controllable, for some pair  $(z^*, b^*) \in S^* \times B$ ,
- or  $(\bar{\alpha}, \bar{\bar{D}})$  is a blocking pair for player 2.

In the latter case, notice that  $(\bar{\alpha}, \bar{\bar{D}})$  is a blocking pair for player 2 *and*  $(\bar{\beta}, \bar{\bar{D}})$  is a blocking pair for player 1. The handling of such a case is more delicate. We use a by-product of Mertens-Neyman value existence proof to conclude that  $\bar{\bar{D}}$  is solvable, or that there is a controllable exit distribution from  $\bar{\bar{D}}$ , based on joint perturbations of the two players. Since we have ruled out non-absorbing solvable sets,  $\bar{\bar{D}}$  is a controlled set.

Therefore, roughly speaking, any blocking pair contains a controlled set, moreover of a simple type.

### 3.2 The difficult step

We deal here with the part of the idea that has been isolated. Let  $D \subseteq S^*$  be a communicating set for  $(\bar{\alpha}, \bar{\beta})$ , and such that both  $(\bar{\alpha}, D)$  and  $(\bar{\beta}, D)$  are blocking pairs. We intend to prove that, given  $D$  is not solvable, there is some controllable exit distribution from  $D$ .<sup>2</sup>

The idea above uses *only* the subharmonic properties of  $v^1$  (resp. of  $v^2$ ) with respect to the kernel  $p(\cdot|z, \bar{\alpha}, \bar{\beta})$  (resp.  $p(\cdot|z, \alpha, \bar{\beta})$ ). Notice that the notion of controlled set involves the payoff function in a quite tangent way, since it appears only through  $v^1$  (or  $v^2$ ), whereas it is of course quite central in the notion of solvable set. Therefore, it is obvious that this part of the proof will use arguments of a completely different nature.

It is convenient, and not restrictive, to assume that  $p(D|z, a, b) = 0$  if  $p(D|z, a, b) < 1$ .

#### 3.2.1 Reminder on $\varepsilon$ -optimal strategies

We start by pointing out a crucial by-product of Mertens-Neyman value existence proof. Let  $(\alpha^\lambda)_{\lambda \leq \lambda_0}$  and  $(\beta^\mu)_{\mu \leq \mu_0}$  be one-parameter families of stationary strategies. Assume that, for every  $z$ ,  $\lambda \leq \lambda_0$ ,  $\mu \leq \mu_0$ ,

$$\lambda r^1(z, \alpha_z^\lambda, \beta_z^\mu) + (1 - \lambda)E[v_\lambda^1|z, \alpha_z^\lambda, \beta_z^\mu] \geq v_\lambda^1(z) \quad (1)$$

---

<sup>2</sup>I know no example of such a set  $B$ . Any such set has to contain at least two ergodic sets for  $(\bar{\alpha}, \bar{\beta})$ .

Then, for every  $\varepsilon > 0$ , there is a strategy  $\sigma$  which after any history plays like some  $\alpha^\lambda$ :  $\sigma(h_n) = \alpha_{z_n}^{\lambda(h_n)}$ , that has the following property. For any strategy  $\tau$  which always plays like some  $\beta^\mu$ , and any initial state  $z$ ,

$$\gamma_n^1(z, \sigma, \tau) \geq v^1(z) - \varepsilon, \text{ for } n \text{ large enough.}$$

Moreover, the same result holds in every subgame: for every  $h_p$ , one has  $E[\bar{g}_n^1 | h_p] \geq v^1(z_p) - \varepsilon$ , provided  $n$  is large (where  $z_p$  is the terminal state of  $h_p$ ).

Observe finally that if the property 1 holds for  $\lambda_0$ , it also holds for the subfamily  $(\alpha^\lambda)_{\lambda \leq \bar{\lambda}_0}$ , for every  $\bar{\lambda}_0 \leq \lambda_0$ . One may therefore require in addition that  $\lambda(h_n)$  be close to zero, for every  $h_n$ .

### 3.2.2 Application

In Mertens-Neyman's proof,  $\alpha^\lambda$  is an optimal strategy of player 1 in the  $\lambda$ -discounted zero-sum game, and there is no restriction on the set of strategies of player 2.

Let us define  $\alpha^\lambda$  as follows. Denote by  $(\bar{\alpha}^\lambda)$  optimal strategies of player 1 in the  $\lambda$ -discounted games such that  $\lim_\lambda \bar{\alpha}^\lambda = \bar{\alpha}$ . For  $z \in D$ , denote  $A_z = \{a \in A, p(D|z, a, \bar{\beta}_z) = 1\}$ , and define  $\alpha_z^\lambda$  as  $\bar{\alpha}_z^\lambda$ , conditioned on  $A_z$ .

Define  $\beta^\mu$  symmetrically. The following facts are easy to prove:

- $\lim_\lambda \alpha^\lambda = \bar{\alpha}$ , and  $\lim_\mu \beta^\mu = \bar{\beta}$ .
- for every  $z$ , the inequality (1) holds, together with the symmetric counterpart for player 2, provided  $\lambda$  and  $\mu$  are close enough to zero.

This has the following consequence. Let  $\varepsilon > 0$ . Provided  $\varepsilon$  is small enough, we can assume that the inequality (1), and its counterpart for player 2, hold for every  $\lambda, \mu < \varepsilon$ , and that moreover  $\|\alpha^\lambda - \bar{\alpha}\|, \|\beta^\mu - \bar{\beta}\| < \varepsilon$ .

Replace the states  $z$  outside  $D$  by absorbing states, with payoff  $v(z) + \varepsilon$ . In this new game, there exists a profile  $(\sigma_\varepsilon, \tau_\varepsilon)$  of strategies, such that:

1. for every history  $h_n$ ,  $\|\sigma_\varepsilon(h_n) - \bar{\alpha}_{z_n}\|, \|\tau_\varepsilon(h_n) - \bar{\beta}_{z_n}\| < \varepsilon$ .
2.  $\gamma_n(z, \sigma_\varepsilon, \tau_\varepsilon) \geq v(z) - \varepsilon$ , for  $n$  large, and the same holds in any subgame.<sup>3</sup>

---

<sup>3</sup>In this inequality,  $\gamma_n(z, \sigma_\varepsilon, \tau_\varepsilon)$  stands for the average payoff in the new game, and  $v(z)$  for the value in the original game. We use here the easy observation that the value of the new game is at least the value of the original game.

Denote by  $p_\varepsilon$  the probability that, starting from  $z$ , under  $(\sigma_\varepsilon, \tau_\varepsilon)$ , the play ever leaves  $D$  (*i.e.*, reaches an absorbing state). We discuss according to the asymptotic behavior of  $p_\varepsilon$ , for  $\varepsilon$  small.

**CASE 1:** there is a sequence  $(\varepsilon_n)$  converging to zero, with  $p_{\varepsilon_n} = 1$  for every  $n$ .

We argue that there is a controllable exit distribution from  $D$ . Let  $\varepsilon$  belong to the sequence, and denote by  $Q_\varepsilon$  the distribution (starting from  $z$ , under  $(\sigma_\varepsilon, \tau_\varepsilon)$ ) of the exit state from  $D$ . By construction,  $p(D|z_n, \sigma_\varepsilon(h_n), \bar{\beta}_{z_n}) = 1 = p(D|z_n, \bar{\alpha}_{z_n}, \tau_\varepsilon(h_n))$ , for every history  $h_n$ . Thus,  $Q_\varepsilon$  belongs to the set which was denoted  $\mathcal{Q}^2(\bar{\alpha}, \bar{\beta})$  in the previous chapter by Solan:  $Q_\varepsilon$  is in the convex hull of the distributions  $p(\cdot|z, a, b)$ , where  $(z, a, b) \in D \times A \times B$ , and  $p(D|z, a, \bar{\beta}_z) = p(D|a, \bar{\alpha}_z, b) = 1$ . Notice that

$$\lim_n \gamma_n(z, \sigma_\varepsilon, \tau_\varepsilon) = Q_\varepsilon v + \varepsilon.$$

Therefore,  $Q_\varepsilon v \geq v(z) - \varepsilon$ . Since  $\mathcal{Q}^2(\bar{\alpha}, \bar{\beta})$  is compact, there is a distribution  $Q \in \mathcal{Q}^2(\bar{\alpha}, \bar{\beta})$  with  $Qv \geq v(z)$ .

Since  $Q$  involves no unilateral exits, and both pairs  $(\bar{\alpha}, D)$  and  $(\bar{\beta}, D)$  are blocking, one concludes that the exit  $Q$  is controllable (with respect to any  $\gamma \geq v$ ).

**CASE 2:** there is a sequence  $(\varepsilon_n)$  converging to zero, with  $p_{\varepsilon_n} < 1$ , for every  $n$ .

We argue that  $D$  is solvable. Denote by  $\mathcal{R}$  the set of ergodic sets for  $(\bar{\alpha}, \bar{\beta})$ , which are included in  $D$ . Choose any  $\varepsilon$  such that  $p_\varepsilon < 1$ . Consider any subgame (*i.e.* any history  $h_p$ ) in which the probability of reaching an absorbing state in finite time is close to 0. Denote by  $\gamma_n(h_p, \sigma_\varepsilon, \tau_\varepsilon)$  the average payoff induced by  $(\sigma_\varepsilon, \tau_\varepsilon)$  in the subgame defined by  $h_p$ . Since  $\sigma_\varepsilon$  and  $\tau_\varepsilon$  always play approximately like  $\bar{\alpha}$  and  $\bar{\beta}$ , the average payoff  $\gamma_n(h_p, \sigma_\varepsilon, \tau_\varepsilon)$  is, for  $n$  large, close to some convex distribution of the payoff vectors  $\gamma(R, \bar{\alpha}, \bar{\beta})$ ,  $R \in \mathcal{R}$ . How close it is depends on  $\varepsilon$ . Since this is true for every  $\varepsilon$ , and since the convex hull of  $\{\gamma(R, \bar{\alpha}, \bar{\beta}), R \in \mathcal{R}\}$  is compact, there is some point  $d$  in this convex hull, with  $d \geq v$ .

Since  $D$  communicates under  $(\bar{\alpha}, \bar{\beta})$ , and both pairs  $(\bar{\alpha}, D)$  and  $(\bar{\beta}, D)$  are blocking, one concludes that  $D$  is solvable.



### 3.3 The algorithm

When one tries to turn the above idea into a proof, one runs into some troubles. One may have to find, *iteratively*, controlled sets within blocking sets. After the first round, one needs to apply the above idea to a reduced game. As mentioned above, the pivotal point in the proof is the above-mentioned by-product of Mertens-Neyman's proof, which enables to relate assumptions on transitions (the fact that  $(\bar{\alpha}, \bar{D})$  and  $(\bar{\beta}, \bar{D})$  are blocking pairs) to properties of *average payoffs* (solvable set). The corresponding would be quite cumbersome.

Therefore; we reverse the argument and proceed in two steps as follows:

**STEP 1 :** let  $(D_1, \dots, D_M)$  be the maximal subsets of  $S^*$  with the following properties, for every  $m$ :

- $v$  is constant on  $D_m$ ;
- $D_m$  communicates for  $(\bar{\alpha}, \bar{\beta})$ ;
- $(\bar{\beta}, D_m)$  is a blocking pair for player 1.

The sets  $(D_m)$  are disjoint. Moreover, the family  $(D_m)$  is uniquely defined (given the choice of  $(\bar{\alpha}, \bar{\beta})$ ).

For each  $m$ , there is a controllable exit distribution from  $D_m$ , based on perturbations of  $(\bar{\alpha}, \bar{\beta})$ , as was shown above. Choose *any* such exit distribution. The following lemma is a simple consequence of the fact that the sets  $D_m$  were chosen maximal.

**Lemma 13** *Consider the reduced game  $\Gamma_{(D_1, \dots, D_M)}$ . There is no set  $D$ , such that  $(\bar{\beta}, D)$  is a reduced blocking pair for player 1.*

**Proof.** Assume such a  $D$  exists. It is easy to check that there is a subset

$\bar{\mathcal{D}}$  of  $D$  such that  $v^2$  is constant on  $\bar{\mathcal{D}}$ <sup>4</sup> and  $(\bar{\beta}, \bar{\mathcal{D}})$  is a reduced blocking pair for player 1. A variation on Lemma 12 shows that there exists a subset  $\bar{D}$  of  $\bar{\mathcal{D}}$  such that  $v$  is constant on  $\bar{D}$  and  $(\bar{\beta}, \bar{D})$  is a reduced blocking pair. It is then clearly a blocking pair for player 1 in the original game  $\Gamma$  (one here identifies  $\bar{D}$  to a subset of  $S^*$ ). Of course,  $\bar{D}$  cannot coincide with a single  $D_m$  (it would not be stable in the reduced game otherwise). Therefore, either

---

<sup>4</sup>Since  $B$  is a subset of  $S^*_{(B_1, \dots, B_M)}$ , we may have to speak of  $v(b_i)$ , where  $b_i$  is the dummy state which stands for  $B_i$ . We set  $v^j(b_i) = \max_{B_i} v^j$ , for  $j = 1, 2$ .

$\overline{D}$  strictly contains some  $D_m$ , or is disjoint from them. In both cases, this contradicts the definition of the sets  $D_1, \dots, D_M$ . ■

**STEP 2:** let  $(C_1, \dots, C_L)$  be subsets of  $S_{(D_1, \dots, D_M)}^*$ , with the following properties, for every  $l$ :

1.  $v^2$  is constant on  $C_l$ ;
2. for some  $\alpha_l$ ,  $C_l$  is communicating for  $(\alpha_l, \overline{\beta})$  and  $(\alpha_l, C_l)$  is a reduced blocking pair for player 2.

We need an equivalent of the maximality assumption used in STEP 1. We add the requirement that, for every  $l \in \{1, \dots, L\}$ ,  $C_l$  is a maximal subset of  $S^* \hat{A} C_1 \cup C_2 \dots \cup C_{l-1}$  with the properties 1 and 2 (and there is no subset of  $S^* \hat{A} C_1 \cup C_2 \dots \cup C_L$  with properties 1 and 2).

The fact that the family  $(C_1, \dots, C_L)$  is not uniquely defined is not important. By STEP 1, none of the pairs  $(C_l, \overline{\beta})$  is a *reduced* blocking pair for player 1. This implies that none of the pairs  $(C_l, \overline{\beta})$  is a blocking pair for player 1. Therefore, the exit distribution  $p(\cdot | z_i, a_i; \overline{\beta})$  from  $C_l$  is controllable, for some  $z_l \in C_l, a \in A$ .

As in STEP 1, the next lemma is a consequence of the maximality property of the sets  $(C_l)$ . Its proof is an adaptation of the proof of Lemma 13.

**Lemma 14** *Consider the reduced game  $(\Gamma_{(D_1, \dots, D_M)})(C_1, \dots, C_L)$ . There is no reduced blocking pair for player 2.*

## 4 Reduced games and recursive games

It remains to prove Proposition 10. The proof relies on the fact that, in the reduced game that we obtained, there is an individually rational, absorbing reply of player 2 to any given  $\alpha$ . This fact relies itself on the idea that the reduction eliminated all blocking pairs for player 2.

**Lemma 15** *The following is true for the reduced game  $\Gamma_C$ . For every  $\alpha$ , there exists  $\beta$  such that the profile  $(\alpha, \beta)$  is absorbing and  $\gamma^2(z, \alpha, \beta) \geq v^2(z)$ .*

We stress once more the fact that  $v^2$  is the value associated to the original game  $\Gamma$ .

**Proof.** choose  $\beta$  such that the support of  $\beta_z$  consists exactly of those  $b$

such that

$$E^{\mathbb{F}} v^2 | z, \alpha_z, b^{\alpha} \geq v^2(z),$$

and notice that  $E_C [v^2 | z, a, b] \geq E [v^2 | z, a, b]$ , for every  $(z, a, b)$ . ■

We now prove Proposition 10.

**Proof.** any  $\varepsilon$ -equilibrium  $(\sigma, \tau)$  of  $\mathbb{P}$  is absorbing. Choose an integer  $N$

such that, under  $(\sigma, \tau)$ , an absorbing state is reached before stage  $N$ , with high probability. Then the profile which plays  $(\sigma, \tau)$  up to stage  $N$ , and punishments strategies afterwards, is an absorbing  $\varepsilon'$ -equilibrium, where  $\varepsilon'$  goes to zero with  $\varepsilon$ .<sup>5</sup> ■

## References

- [1] Mertens, J.F. and A. Neyman (1981). Stochastic games, *International Journal of Game Theory*, **10**, 53-66.
- [2] Solan, E. (1999). General Tools - Perturbations of Markov Chains, *this book*.
- [3] Thuijsman, F. and O.J. Vrieze (1991). Easy initial states in stochastic games, in *Stochastic games and related topics*, eds. T.S. Ferguson, T.E.S. Raghavan, T. Parthasarathy and O.K. Vrieze, Kluwer Academic Publishers.
- [4] Vieille, N. (1997). Equilibrium in two-player stochastic games I: a reduction, CEREMADE D.P. 9745.
- [5] Vieille, N. (1997). Equilibrium in two-player stochastic games II: the case of recursive games, CEREMADE D.P. 9747.
- [6] Vieille, N (1998). Solvable states in  $N$ -player games, CEREMADE D.P. 9812, to appear in *SIAM Journal of Control and Optimization*.

---

<sup>5</sup>To properly conclude, one needs to know how the value of  $\mathbb{P}$  is related to the value of the reduced game. I will not elaborate on this point.